V. Sulimov[1,2]
D. Kutov[1,2],
A. Sulimov[1,2]

DOCKING PROBLEMS: PROTEIN-LIGAND LOW ENERGY MINIMA ANALYSIS

[1] Dimonta, Ltd, Nagornaya Street 15, bldg 8, Moscow 117186, Russia;
[2] Research Computer Center, Moscow State University, Leninskie Gory 1, bldg 4, Moscow 119992, Russia;

Vladimir.sulimov@gmail.com

Search of molecules-inhibitors of a given target protein is the key stage of the new drug development. Molecular modeling by docking and molecular dynamics programs should increase effectiveness of new inhibitors development. The reliable prediction is defined by the accuracy of these programs. Molecular dynamics and docking have many common features limiting their accuracy. The lecture will be devoted to main problems limiting accuracy of docking programs and how to solve them. Docking programs perform positioning of a compound (a ligand) in the active site of the target protein. Computed poses of the ligand are used for the calculation of the protein-ligand binding free energy which is directly connected with the inhibition constant. The protein-ligand binding energy ΔG_{bind} is calculated as the difference between the free energy of the protein-ligand complex G_{PL} and the sum of free energies of the unbound protein G_P and the ligand G_L:

$$\Delta G_{bind} = G_{PL} - G_P - G_L$$

Free energies of the protein, the ligand and their complex are described by respective energy landscapes and they can be calculated through the configuration integrals over the respective phase space. In the thermodynamic equilibrium the molecular system occupies its low energy minima. The configuration integral will come to the sum of configuration integrals over the separate low energy minima if these minima are separated by sufficiently high energy barriers [1]. So, the docking accuracy is defined by the completeness of the low energy minima spectra of the molecular systems and by the accuracy of the configuration integral calculation in each of these minima. Main simplifications of many existing docking programs is the rigid protein approximation and the use of the preliminary calculated grid of potentials of ligand probe atoms interactions with the protein (the grid approximation) restricting performance of docking and worsening its accuracy. In this study we describe several docking programs including “classical” docking programs on the base of grid approximation for virtual screening of large ligand databases, e.g. SOL [2], and novel direct generalized docking programs FLM [1], SOL-T [3] and SOL-P [4]. The latter makes it possible to reject the rigid protein as well as the grid approximations, to take into account many proteins’ degrees of freedom and to increase the docking accuracy. The docking programs are based on the paradigm which assumes that the ligand binding pose in the active site of the target protein corresponds to the global minimum of the protein-ligand energy or is near it and the docking problem is reduced to the global optimization problem on the multi-dimensional protein-ligand energy surface. Detailed analysis of low energy local minima demonstrates feasibility of the docking paradigm. Genetic and tensor train global optimization algorithms are brifly described as a base for docking programs. The latter make it possible to perform successful docking in the conformation space of 157 degrees of freedom: a flexible ligand and several dozen of moveable protein atoms. Mobility of protein atoms increases docking positioning accuracy. Importance of solvent accounting in the docking procedure is demonstrated. Quasi-docking procedure for testing applicability of different force fields and quantum-chemical methods for docking is presented. The important role of multi-processors supercomputer calculations in docking is demonstrated. Future perspectives the docking development are described.

1. Oferkin I. et al. *Advances in Bioinformatics*, 2015, **2015**, Article ID 126858, 12 pages.
2. Sulimov A. et al. *Journal of Chemical Information and Modeling*, 2013, **53**: 1946-1956.
3. Oferkin I. et al. *Bulletin of the South Ural State University, Ser. Mathematical Modelling, Programming & Computer Software*, 2015, **8**: 83–99.
4. Sulimov A. et al. *Computational and Structural Biotechnology Journal*, 2017, **15**: 275-285.

The work was financially supported by the Russian Science Foundation, Agreement no. 15-11-00025